

II. GRAMMAR DAN BAHASA

Konsep Dasar

1. Dalam pembicaraan grammar, anggota alfabet dinamakan simbol terminal atau token.
2. Kalimat adalah deretan hingga simbol-simbol terminal.
3. Bahasa adalah himpunan kalimat-kalimat. Anggota bahasa bisa tak hingga kalimat.
4. Simbol-simbol berikut adalah simbol terminal :
 - huruf kecil awal alfabet, misalnya : a, b, c
 - simbol operator, misalnya : +, -, dan ·
 - simbol tanda baca, misalnya : (,), dan ;
 - string yang tercetak tebal, misalnya : **if**, **then**, dan **else**.
5. Simbol-simbol berikut adalah simbol non terminal :
 - huruf besar awal alfabet, misalnya : A, B, C
 - huruf S sebagai simbol awal
 - string yang tercetak miring, misalnya : *expr* dan *stmt*.
6. Huruf besar akhir alfabet melambangkan simbol terminal atau non terminal, misalnya : X, Y, Z.
7. Huruf kecil akhir alfabet melambangkan string yang tersusun atas simbol-simbol terminal, misalnya : x, y, z.
8. Huruf yunani melambangkan string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya, misalnya : α , β , dan γ .
9. Sebuah produksi dilambangkan sebagai $\alpha \rightarrow \beta$, artinya : dalam sebuah derivasi dapat dilakukan penggantian simbol α dengan simbol β .
10. Simbol α dalam produksi berbentuk $\alpha \rightarrow \beta$ disebut ruas kiri produksi sedangkan simbol β disebut ruas kanan produksi.
11. Derivasi adalah proses pembentukan sebuah kalimat atau sentensial. Sebuah derivasi dilambangkan sebagai : $\alpha \Rightarrow \beta$.
12. Sentensial adalah string yang tersusun atas simbol-simbol terminal atau simbol-simbol non terminal atau campuran keduanya.
13. Kalimat adalah string yang tersusun atas simbol-simbol terminal. Jelaslah bahwa kalimat adalah kasus khusus dari sentensial.
14. Pengertian terminal berasal dari kata *terminate* (berakhir), maksudnya derivasi berakhir jika sentensial yang dihasilkan adalah sebuah kalimat (yang tersusun atas simbol-simbol terminal itu).
15. Pengertian non terminal berasal dari kata *not terminate* (belum/tidak berakhir), maksudnya derivasi belum/tidak berakhir jika sentensial yang dihasilkan mengandung simbol non terminal.

Grammar dan Klasifikasi Chomsky

Grammar G didefinisikan sebagai pasangan 4 tuple : V_T, V_N, S , dan Q , dan dituliskan sebagai $G(V_T, V_N, S, Q)$, dimana :

V_T : himpunan simbol-simbol terminal (atau himpunan token -token, atau alfabet)

V_N : himpunan simbol-simbol non terminal

$S \in V_N$: simbol awal (atau simbol start)

Q : himpunan produksi

Berdasarkan komposisi bentuk ruas kiri dan ruas kanan produksinya ($\alpha \rightarrow \beta$), Noam Chomsky mengklasifikasikan 4 tipe grammar :

1. Grammar tipe ke-0 : Unrestricted Grammar
(UG) Ciri : $\alpha, \beta \in (V_T \mid V_N)^*$, $\alpha > 0$
2. Grammar tipe ke-1 : Context Sensitive Grammar (CSG)
Ciri : $\alpha, \beta \in (V_T \mid V_N)^*$, $0 < \alpha \leq \beta$
3. Grammar tipe ke-2 : Context Free Grammar
(CFG) Ciri : $\alpha \in V_N$, $\beta \in (V_T \mid V_N)^*$
4. Grammar tipe ke-3 : Regular Grammar (RG)
Ciri : $\alpha \in V_N$, $\beta \in \{V_T, V_T V_N\}$ atau $\alpha \in V_N$, $\beta \in \{V_T, V_N V_T\}$
Meningat ketentuan simbol-simbol (hal. 3 no. 4 dan 5), ciri-ciri RG sering dituliskan sebagai : $\alpha \in V_N$, $\beta \in \{a, bC\}$ atau $\alpha \in V_N$, $\beta \in \{a, Bc\}$

Tipe sebuah grammar (atau bahasa) ditentukan dengan aturan sebagai berikut :

A language is said to be type- i ($i = 0, 1, 2, 3$) language if it can be specified by a type- i grammar but can't be specified any type- $(i+1)$ grammar.

Contoh Analisa Penentuan Type Grammar

1. Grammar G_1 dengan $Q_1 = \{S \rightarrow aB, B \rightarrow bB, B \rightarrow b\}$. Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_1 kemungkinan tipe CFG atau RG. Selanjutnya karena semua ruas kanannya terdiri dari sebuah V_T atau string $V_T V_N$ maka G_1 adalah RG.
2. Grammar G_2 dengan $Q_2 = \{S \rightarrow Ba, B \rightarrow Bb, B \rightarrow b\}$. Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_2 kemungkinan tipe CFG atau RG. Selanjutnya karena semua ruas kanannya terdiri dari sebuah V_T atau string $V_N V_T$ maka G_2 adalah RG.
3. Grammar G_3 dengan $Q_3 = \{S \rightarrow Ba, B \rightarrow bB, B \rightarrow b\}$. Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_3 kemungkinan tipe CFG atau RG. Selanjutnya karena ruas kanannya mengandung string $V_T V_N$ (yaitu bB) dan juga string $V_N V_T$ (Ba) maka G_3 bukan RG, dengan kata lain G_3 adalah CFG.
4. Grammar G_4 dengan $Q_4 = \{S \rightarrow aAb, B \rightarrow aB\}$. Ruas kiri semua produksinya terdiri dari sebuah V_N maka G_4 kemungkinan tipe CFG atau RG. Selanjutnya karena ruas kanannya mengandung string yang panjangnya lebih dari 2 (yaitu aAb) maka G_4 bukan RG, dengan kata lain G_4 adalah CFG.
5. Grammar G_5 dengan $Q_5 = \{S \rightarrow aA, S \rightarrow aB, aAb \rightarrow aBCb\}$. Ruas kirinya mengandung string yang panjangnya lebih dari 1 (yaitu aAb) maka G_5 kemungkinan tipe CSG atau UG. Selanjutnya karena semua ruas kirinya lebih pendek atau sama dengan ruas kananya maka G_5 adalah CSG.
6. Grammar G_6 dengan $Q_6 = \{aS \rightarrow ab, SAc \rightarrow bc\}$. Ruas kirinya mengandung string yang panjangnya lebih dari 1 maka G_6 kemungkinan tipe CSG atau UG. Selanjutnya karena terdapat ruas kirinya yang lebih panjang daripada ruas kananya (yaitu SAc) maka G_6 adalah UG.

Derivasi Kalimat dan Penentuan Bahasa

Tentukan bahasa dari masing-masing grammar berikut :

1. G_1 dengan $Q_1 = \{1. S \rightarrow aAa, 2. A \rightarrow aAa, 3. A \rightarrow b\}$.

Jawab :

Derivasi kalimat terpendek :

$$S \Rightarrow aAa \quad (1)$$

$$\Rightarrow aba \quad (3)$$

Derivasi kalimat umum :

$$S \Rightarrow aAa \quad (1)$$

$$\Rightarrow aaAaa \quad (2)$$

...

$$\Rightarrow a^n Aa^n \quad (2)$$

$$\Rightarrow a^n ba^n \quad (3)$$

Dari pola kedua kalimat disimpulkan : $L_1(G_1) = \{ a^n ba^n \mid n \geq 1 \}$

2. G_2 dengan $Q_2 = \{1. S \rightarrow aS, 2. S \rightarrow aB, 3. B \rightarrow bC, 4. C \rightarrow aC, 5. C \rightarrow a\}$.

Jawab :

Derivasi kalimat terpendek :

$$S \Rightarrow aB \quad (2)$$

$$\Rightarrow abC \quad (3)$$

$$\Rightarrow aba \quad (5)$$

Derivasi kalimat umum :

$$S \Rightarrow aS \quad (1)$$

...

$$\Rightarrow a^{n-1} S \quad (1)$$

$$\Rightarrow a^n B \quad (2)$$

$$\Rightarrow a^n bC \quad (3)$$

$$\Rightarrow a^n baC \quad (4)$$

...

$$\Rightarrow a^n ba^{m-1} C \quad (4)$$

$$\Rightarrow a^n ba^m \quad (5)$$

Dari pola kedua kalimat disimpulkan : $L_2(G_2) = \{ a^n ba^m \mid n \geq 1, m \geq 1 \}$

3. G_3 dengan $Q_3 = \{1. S \rightarrow aSBC, 2. S \rightarrow abC, 3. bB \rightarrow bb, 4. bC \rightarrow bc, 5. CB \rightarrow BC, 6. cC \rightarrow cc\}$.

Jawab :

Derivasi kalimat terpendek 1:

$$S \Rightarrow abC \quad (2)$$

$$\Rightarrow abc \quad (4)$$

Derivasi kalimat terpendek 2 :

$$S \Rightarrow aSBC \quad (1)$$

$$\Rightarrow aabCBC \quad (2)$$

$$\Rightarrow aabBCC \quad (5)$$

$$\Rightarrow aabbCC \quad (3)$$

$$\Rightarrow aabbcC \quad (4)$$

$$\Rightarrow aabbcc \quad (6)$$

Derivasi kalimat terpendek 3 :

$$S \Rightarrow aSBC \quad (1)$$

$$\Rightarrow aaSBCBC \quad (1)$$

$$\Rightarrow aaabCBCBC \quad (2)$$

$$\Rightarrow aaabBCCBC \quad (5)$$

$$\Rightarrow aaabBCBCC \quad (5)$$

$$\Rightarrow aaabBBCCC \quad (5)$$

$$\Rightarrow aaabbBCCC \quad (3)$$

$$\Rightarrow aaabbbCCC \quad (3)$$

$$\Rightarrow aaabbbcCC \quad (4)$$

$$\Rightarrow aaabbbccC \quad (6)$$

$$\Rightarrow aaabbbccc \quad (6)$$

Dari pola ketiga kalimat disimpulkan : $L_3(G_3) = \{ a^n b^n c^n \mid n \geq 1 \}$

